

Cross-modal discrimination of human gender by domestic dogs

Article (Unspecified)

Ratcliffe, V F, McComb, K and Reby, D (2014) Cross-modal discrimination of human gender by domestic dogs. *Animal Behaviour*, 91. pp. 127-135. ISSN 0003-3472

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/50509/>

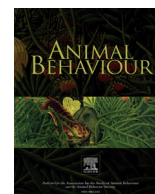
This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Cross-modal discrimination of human gender by domestic dogs



V. F. Ratcliffe, K. McComb, D. Reby*

School of Psychology, University of Sussex, Falmer, U.K.

ARTICLE INFO

Article history:

Received 19 September 2013

Initial acceptance 31 October 2013

Final acceptance 14 February 2014

Available online 14 April 2014

MS. number: 13-00785R

Keywords:

categorization
cross-modal association
domestic dog
gender
social experience
vocal communication

We spontaneously categorize people as male or female, and when hearing a human voice we expect to see an appropriate sex-matched visual image. The extent to which domesticated species, which share our social environment, spontaneously develop such categorization abilities remains underinvestigated. Here we used a cross-modal preferential looking design to determine whether domestic dogs, *Canis familiaris*, spontaneously attribute an unfamiliar voice to a person of the corresponding sex. Fifty-one dogs were played a prerecorded male or female voice in the presence of a man and a woman. The responses were scored as correct or incorrect from both the direction of the first look and the total gaze duration towards each person after the voice presentation. Dogs living with one adult, or one man and one woman, performed significantly below chance as more (71%) of these dogs looked towards the incorrect person first. However, dogs living with more than two adults (including at least one man and one woman) performed significantly better, and significantly more (80%) of these dogs looked at the correct person for longer than they looked at the incorrect person. This suggests that while all of the dogs had spontaneously learnt to categorize human gender across sensory modalities, this ability was expressed differently depending on their social experience with humans. Dogs with greater experience, through regular exposure to multiple male and female human exemplars, responded by orienting towards the correct person, while those with more limited experience avoided looking towards the correct person. We discuss the importance of experience in determining the way that individuals spontaneously form and express categorization abilities.

© 2014 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Categorization is a key cognitive mechanism that determines how we perceive and process sensory information. As well as simplifying processing requirements (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), organizing stimuli into categories allows general inferences to be made and applied to new category members. Humans readily form complex, hierarchical categories representing their environment, using language to create specific referents that can coordinate categories between individuals (see Steels & Belpaeme, 2005 for a review). Currently, only a small number of studies have explored spontaneous category formation in other species, focusing on nonhuman primates (e.g. Murai, Tomonaga, Kamegai, Terazawa, & Yamaguchi, 2004; Murai et al., 2005). Comparative investigations into spontaneous category formation in nonhuman animals are therefore necessary to determine the functional relevance of this cognitive process in a broader range of species.

Domestic dogs provide an interesting model species to compare natural category formation in animals and humans. Dogs have

shared the same environment as humans for at least 15 000 years (Savolainen, Zhang, Luo, Lundberg, & Leitner, 2002), during which time they are likely to have undergone selection promoting specific sociocognitive abilities that allow effective cooperation and communication between the two species (Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare, Brown, Williamson, & Tomasello, 2002). Added to this evolutionary predisposition is the effect of experience, as many dogs are extensively socialized with people, often sharing the same living habitat from an early age. As the human environment has become functionally relevant to dogs, this species may be expected to form spontaneous categories that are directly comparable with human categories.

It has already been established that, with training, dogs show equivalent categorization abilities to other mammals and birds. They are able to discriminate between 'dog' and 'nondog' sounds (Heffner, 1975), images of dogs and landscapes (Range, Aust, Steurer, & Huber, 2008) and images of dogs and other species (Autier-Dérain, Deputte, Chalvet-Monfray, Coulon, & Mounier, 2013), correctly generalizing their responses to novel stimuli. Spontaneous, ecologically relevant category formation is also evident in the dog's ability to form cross-modal perceptual associations when responding to familiar people. Using an expectancy

* Correspondence: D. Reby, School of Psychology, University of Sussex, Falmer BN1 9QH, U.K.

E-mail address: d.reby@sussex.ac.uk (D. Reby).

violation paradigm, Adachi, Kuwahata, and Fujita (2007) presented dogs with a photograph of either their owner's or a stranger's face after playing back one of their voices. Dogs looked for longer when the face did not match the preceding voice than when the stimuli did match, suggesting that dogs can use cross-modal associative categories when responding to familiar humans. This form of categorical perception is likely to be expressed naturally by dogs as the need to identify familiar humans has a clear function in recognizing important social partners and care providers. The spontaneous use of an interspecific category representing familiar humans leads to the possibility that it may also be relevant for dogs to form categories about unfamiliar humans, which would allow direct comparisons with our own categories.

One of the predominant ways that we categorize unfamiliar people is by their gender, primarily by associating visual and vocal cues. Because human faces are sexually dimorphic (Burton, Bruce, & Dench, 1993), differing in both shape and texture (Hill, Bruce, & Akamatsu, 1995), face gender classification in adults is close to 100% accuracy (O'Toole et al., 1998). Sexual dimorphism also leads to differences in the vocal tract anatomy of adult men and women. The larger adult male larynx results in a difference of approximately 80 Hz in fundamental frequency (F0) between the voices of adult men and women, with mean values at around 120 Hz and 200 Hz, respectively (Titze, 2000). Additionally, adult men have a disproportionately longer vocal tract than women (Vorperian et al., 2009), causing lower first formant (F1) values and formant dispersions approximately 15–20% lower than in women (Fant, 1960; Goldstein, 1980). The relative F0 and formant values classify the gender of adult voices at 98.8% accuracy (Bachorowski & Owren, 1999). The presence of both visual and vocal gender cues enables cross-modal perceptual matching of voices to individuals from an early age (Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991).

Dogs are also likely to be able to perceive these gender differences in the human voice, as they attend to variation in formants to determine size information in conspecific vocalizations (Taylor, Reby, & McComb, 2011) and can be trained to discriminate between average male and female F0 differences in human vowel sounds (Baru, 1975). Gender-specific behavioural differences in the way humans interact with dogs have been identified (e.g. Prato-Previde, Fallani, & Valsecchi, 2005) which could have created the need for dogs to categorize human gender in order to adjust their responses appropriately. In support of this, shelter-housed dogs petted by women show more relaxed behaviour and lower cortisol levels than those petted by men (Hennessy, Williams, Miller, Douglas, & Voith, 1998), and are more likely to direct defensive aggressive behaviour towards men than women (Lore & Eisenberg 1986; Wells & Hepper, 1999). Although the specific cues to which the dogs were responding cannot be determined from these studies, they do suggest that categorically assessing human gender could be functionally relevant for dogs, influencing their reaction to the individual person. Therefore the ability to perceive and associate different sensory cues to human gender as categorically equivalent would be a useful ability.

To determine whether dogs do categorize human gender using different sensory cues, we tested whether they associate voices with unfamiliar people using gender cues in a cross-modal preferential looking paradigm, in which subjects were required spontaneously to match voices to people by their gender. In our study, a man and a woman stood either side of a loudspeaker from which a voice recording of a different person was played. Dogs were positioned facing the centre line, and their visual orientation to the person matching the gender of the voice and the nonmatching person were recorded. If dogs spontaneously combine vocal and visual cues to identify human gender cross-modally, it was predicted that they would look first, and for a longer duration, at the

person of the same gender as the voice. The potential effect of social factors on performance was also investigated, as well as possible mechanisms involved in such variation.

METHODS

Subjects

A total of 51 adult dogs of 17 different breeds were recruited when their owners responded to advertisements in the East Sussex area. Ages ranged from 7 months to 11 years old (mean + SD = 5.03 + 3.17 years), including 26 males and 25 females. The selection criteria for subject animals were that they had to be healthy adults (older than 6 months) with no known sight or hearing problems and no known aggression towards people. Subjects and their owners were naïve to the experimental set-up and had not participated in any previous vocal communication or behavioural research.

Playback Acquisition

Nine men and nine women, aged between 20 and 52 years (mean + SD = 30.94 + 9.75 years), were audio recorded after being instructed to pronounce the following phrases as if speaking to a dog in a positive voice: 'Hey!', 'Come on then', 'Good dog!', 'What's this?'. Each speaker pronounced each phrase once. All recordings were made using a Zoom H4N Handy Recorder in a soundproof booth. The sampling frequency was set at 44 100 Hz, with a 32-bit sampling rate, for each recording. The vocal parameters of the recordings were then checked for a bimodal distribution according to gender using PRAAT v5.0.3 (<http://www.fon.hum.uva.nl/praat/>). The four phrases were analysed together as a single audio file. The mean, minimum and maximum F0 values were calculated using the PRAAT autocorrelation algorithm 'to Pitch (ac)' which estimates the F0 contour across the utterance. The mean F0 for the male voices was between 142.00 Hz and 193.48 Hz (mean + SD = 166.80 + 17.64 Hz), while the mean F0 for the female voices was between 251.13 Hz and 405.99 Hz (mean + SD = 323.26 + 61.22 Hz). The F0 ranges (maximum F0–minimum F0) for the male voices were 109.70–164.73 Hz (mean + SD = 154.16 + 48.32 Hz), while the female F0 ranges were 269.41–528.01 Hz (mean + SD = 350.20 + 86.35 Hz). The formant dispersion (ΔF) was calculated using the PRAAT Linear Predictive Coding 'Burg' algorithm, which estimates the centre frequencies of the first four formants across the utterance. These values were then used to calculate the average spacing between the formants. The male ΔF s were between 927.98 Hz and 1120.60 Hz (mean + SD = 1029.15 + 71.39 Hz), while the female ΔF s were between 1140.60 Hz and 1241.00 Hz (mean + SD = 1215.56 + 45.20 Hz). All of the recordings were normalized to –1.0 dB maximum amplitude in Audacity 2.0.0 (<http://audacity.sourceforge.net>).

Experimental Set-up

Experiments were carried out between June and September 2012 at two indoor test locations in the East Sussex, U.K., area (The Dog Hut in Barcombe and Hamsey Riding School in Lewes). A cross-modal preferential looking paradigm was used. The design was developed on the basis of pilot trials conducted in April and May 2012 on 20 subjects, who did not take part in the final study trials. The original piloted study included a sequence of six trials per subject; however, we found that habituation to the procedure led to a reduction in responses after the first trial. Therefore in the full study each dog took part in only one trial.

An Anchor LIB-6000H Liberty loudspeaker (frequency response: 60 Hz–15 kHz) was mounted onto a 130 cm tall stand and

disguised using brown material. The speaker was placed 500 cm in front of a designated subject area (150 cm²) where subjects could be positioned, with a chair for their owner and a black screen placed directly behind them. A SONY DCR-HC51 Handycam video camera was mounted onto a tripod 30 cm from the floor, and positioned directly in front of the loudspeaker, facing and zoomed in towards the subject. Fluorescent coloured rope was used to delineate a centre line between the subject and the loudspeaker and was clearly visible on the videos, providing a visual determinant of left and right during video coding. A second video camera was mounted on a tripod 100 cm from the floor, and placed behind the subject area, facing the loudspeaker. This was to monitor the subject's field of view.

Two assistants, a man and a woman chosen from a pool of 10 people, stood facing the subjects with their nearest foot 150 cm either side of the centre of the loudspeaker (Fig. 1). Each assistant remained stationary with one hand covering their mouth and gazed straight ahead with a neutral facial expression. The assistants did not make eye contact or interact with the dogs in any way throughout the study. The side on which the male and female assistants stood was counterbalanced across subjects. The two assistants were chosen on a pseudorandomized basis from a pool of five men and five women, providing a range in physical attributes including age and hairstyle. The heights of the assistants were bimodally distributed by gender (male range 180.34–190.50 cm, mean + SD = 183.27 + 4.15 cm; female range 153.00–171.00 cm, mean + SD = 162.53 + 7.89 cm).

The visual acuity of dogs is typically given to be around 20/75 using the Snellen fraction (Miller & Murphy, 1995), and therefore their visual perception of objects is less detailed than human visual perception at the same distance. However, using a similar paradigm to the current study, Faragó et al. (2010) demonstrated that dogs were able to discriminate between size-matched images of cats and dogs (approximately 30 cm in height) from a distance of 5 m.

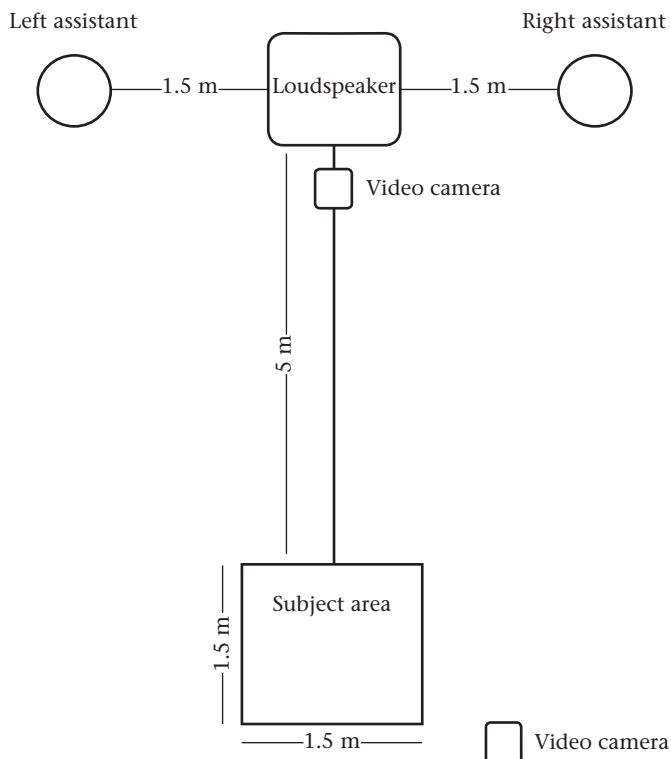


Figure 1. Experimental set-up with distances between the subject, loudspeaker and assistants.

Therefore it was expected that in the present study, as well as acquiring possible olfactory information, dogs would also be able to obtain sufficient visual information from the humans to discriminate gender-related information from the same distance.

Procedure

Subjects were held loosely by their lead throughout the experiment and were always handled by their owner. The owners were naïve to the purpose of the experiment and were instructed on entering the test room to allow their dog to familiarize itself with the area, including walking in front of both the assistants and the loudspeaker. This was aimed at giving the dog access to olfactory information from the assistants. The owner was then instructed to sit on the chair provided and place their dog in front of them. The owner was asked to remain silent and still, and avoid interacting with their dog unless necessary to keep their dog inside the subject area. Neither the owner nor the experimenter was in the subject's field of vision during the experiment, both to minimize unconscious cueing and to prevent dogs matching the playback voice to unintended targets. To ensure that the owner and assistants were not giving any unintentional cues when they heard the playback voice, half of the tests were conducted with both of the assistants and the owner listening to music from handheld MP3 players, which masked the sound of the playback voices.

After settling in their position the subjects were given 10 s for further visual familiarization with the assistants. This was followed with 10 s of silence, followed by the presentation of a single playback voice and a further 10 s of silence. The playback voice consisted of one person saying the four phrases outlined above, in the same order, with a 500 ms interval between each phrase. The recording was played at 65 dB (+/−5 dB), measured by a N05CC Digital Mini Sound Level Meter. The playback exemplar was chosen from the pool of 18 voice recordings in a pseudorandomized order across subjects, so that half of the subjects heard a female voice and half heard a male voice. The individual playback voices, the gender of the voice and the side on which the male and female assistants stood were counterbalanced across subjects.

Collection and Coding of Dog Contextual Information

Following the experiment the owners were asked to fill out a short questionnaire about their dog. This included questions about their dog's breed, age, sex, reproductive status, the number of adult men and women living with the dog (household composition), the average amount of time the dog spent with people per day, behavioural characteristics around unfamiliar men and women and their dog's origination (private or animal shelter).

To explore potential differences between the subjects' performance depending on their human social environment, household composition was also coded as a categorical variable (HC Group) with two levels: subjects living with either one adult person or one man and one woman (HC Group 1, $N = 35$) and subjects living with between three and five adults, including at least one man and one woman (HC Group 2, $N = 15$). The number of men and women living in the household was evenly balanced for the majority of dogs in HC Group 2, as 10 of the subjects lived with two men and two women.

Ethical Note

The dogs were privately owned and handled by their owner throughout the study, which was designed to replicate a routine interaction between a dog and an unfamiliar person. The study complied with the internal University of Sussex regulations on the use of animals and was approved by the University of Sussex Ethical

Review Committee (Approval number: ERC/33/3). Approval to record human voices to be used as stimuli was also obtained from the University of Sussex Life Sciences & Psychology Cluster based Research Ethics Committee (Approval number: DRVR0312).

Behavioural Measures and Coding

Videos were coded in 100 ms intervals using the digital video analysis software Sportscode Gamebreaker version 7.5.5 (Sportstec, Warriewood, NSW, Australia).

The dogs' responses were measured during the 10 s of silence immediately before the playback voice presentation (preplayback) and 10 ms after the onset of the playback voice for a total duration of 15 s until the end of the trial (trial duration was determined from the maximum response duration during the pilot study). The latency, duration and direction of each look (towards each of the assistants, the loudspeaker and away) were recorded in milliseconds.

A look was defined as being at either of the assistants if the dog's head was directed between 15° and 25° from the centre line (delineated by the fluorescent rope), and was recorded as being at the loudspeaker if the dog's head was directed between 0° and 5° from the centre. Finally, a look was recorded as away if it was directed between 6° and 14° or over 26° from the centre point. The orientation of the dog's head was taken by drawing a line from the centre of the top of the dog's forehead to the centre of its nose (Fig. 2). A protractor was placed along the centre line of each video to determine these angles. Although this method does not give an absolute measure of the visual orientation, it does provide a repeatable and standardized index of orientation across the subjects.

From these analyses we derived the following gaze response variables: the total gaze duration (total time looking) and response latency (time to first look) towards each of the assistants, and the total time spent looking elsewhere (at the loudspeaker or away). To characterize the dogs' ability to match the voice gender to the correct assistant, we attributed two separate binary 'correct matching scores' of correct or incorrect using both of the gaze response variables directed towards the assistants (total gaze duration and response latency) after the playback voice presentation. The direction of the first look was scored as correct if the subject looked towards the correct assistant before the incorrect assistant (First look correct score), while the total duration was scored as correct if the subject looked at the correct assistant longer than they looked at the incorrect assistant in total (Total gaze correct score).

To investigate mechanisms underlying gaze responses, we also recorded the occurrence of appeasement signals, which dogs use to reduce the potential of conflict during social interactions. These behaviours are often produced when a dog feels anxious or threatened, and are likely to co-occur with turning away or gaze aversion (Rugaas, 2005). The frequency of occurrence of each of the following appeasement behaviours was monitored: licking the muzzle, yawning, trembling, scratching, sniffing the floor or attempting to move away. The frequencies for each of the behaviours were then summed to provide the total frequency across all behaviours. Any vocalizations made were also recorded and each call type (bark, whine, growl or howl) was scored using the following scale (0: no occurrence; 1: produced less than five times; 2: produced five or more times). The scores for each call type were then added to the total frequency, to provide an overall appeasement behaviour (AB) score for each subject. Two separate AB scores were given to each subject: one score for the time period before (preplayback) and one score after the playback voice presentation.

The videos were coded in blind order by V.R. A research assistant second-coded 84% of the videos, which resulted in a strong

interobserver correlation for both the response latency (Pearson: $r_{40} = 0.80$, $P < 0.001$) and the total gaze duration towards each person (Pearson: $r_{40} = 0.87$, $P < 0.001$). A research assistant also second-coded the appeasement signal scores in 66% of the videos, again resulting in a strong interobserver correlation for the preplayback AB scores (Spearman: $r = 0.78$, $N = 33$, $P < 0.001$) and AB scores after the playback presentation (Spearman: $r = 0.96$, $N = 33$, $P < 0.001$).

One male subject was excluded from subsequent analyses as he did not look at either assistant after hearing the playback voice, giving a total of 50 subjects in the statistical analyses.

Statistical Analysis

Preplayback behaviour

To determine whether the subjects showed any orientation biases prior to the presentation of the auditory stimuli, we ran a mixed factorial ANOVA to test for effects of the gender of the assistant, and/or the side on which they were stood, on the total gaze duration towards each assistant during the first 10 s before the playback voice was presented.

Playback response scores

Binomial probability tests were carried out on the Total gaze and First look correct scores to determine whether the proportion of correct responses differed significantly from the expected 50% chance level. To test the effect of potentially relevant independent variables (IVs) on correct scores, we ran binary logistic regressions with subject's sex, side of the correct person, gender of the playback voice, test location and use of headphones as categorical predictors and subject's age, number of adult people living with the subject (household composition), average number of hours the subject spent with people per day, and difference in height (cm) between the male and female assistants as continuous predictors. Interactions were included between each of the variables related to the subject with variables related to the experimental procedure (gender of the playback voice, side of the correct person and assistant height difference). A forwards stepwise method with a likelihood ratio statistic was used to construct the model by including significant IVs. The same binary logistic regression analyses were also repeated using the categorically coded version of household composition (HC Group). Planned comparisons were then conducted on the significant IVs, using HC Group to identify differences at a group level.

To investigate the potential mechanisms underlying differences in behavioural responses, a mixed factorial ANOVA was performed to test whether the anxiety levels displayed by the subjects (measured by the appeasement behaviour (AB) score) differed before or after the presentation of the playback voice (Time), between HC Groups 1 and 2 (HC Group), or depended on the side of the correct assistant (Side). Finally, to test whether gaze aversion, a common appeasement signal produced by dogs (Rugaas, 2005), could explain observed differences in performance, we tested whether dogs from smaller households (HC Group 1) spent more time looking away from either the correct and/or both assistants than dogs from larger households (HC Group 2).

All analyses were conducted using SPSS version 19 (SPSS Inc., Chicago, IL, U.S.A.).

RESULTS

Preplayback Gazing Behaviour

Analysis of dogs' behaviour during the 10 s prior to playback presentation showed that there was no significant difference in the total gaze duration towards the assistant stood on the

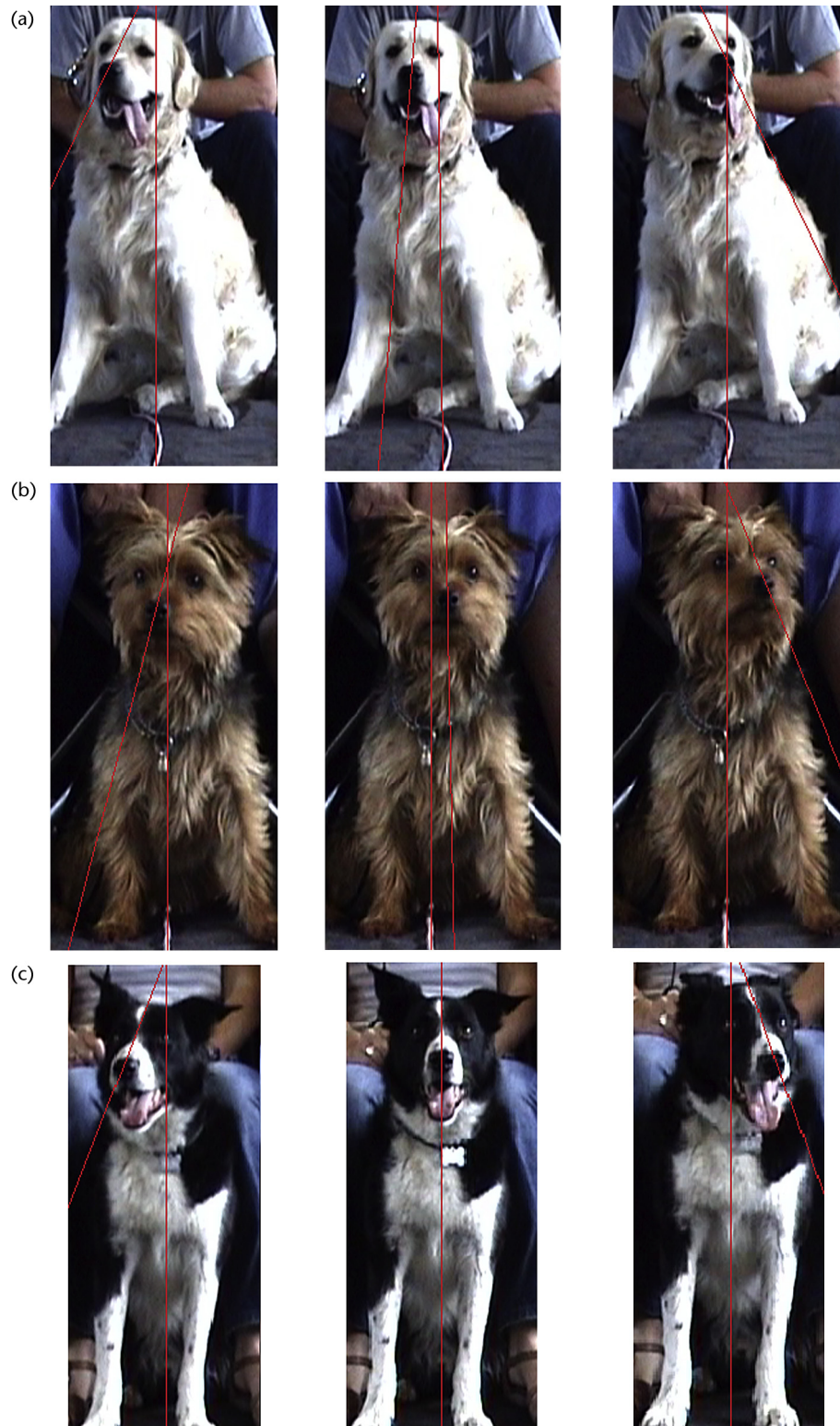


Figure 2. Example frames showing the video analysis coding of the gaze orientation of three subjects. The lines illustrate the angle of the dog's head in relation to the centre line. (a) Golden retriever oriented towards, from left to right frames, the person stood on their right (24°), the loudspeaker (4°) and the person stood on their left (24°). (b) Yorkshire terrier oriented towards the person stood on their right (15°), the loudspeaker (1°) and the person stood on their left (22°). (c) Border collie oriented towards the person stood on their right (21°), the loudspeaker (0°) and the person stood on their left (20°).

subject's left (mean + SE = $997.25 + 259.49$ ms, $N = 50$) or right (mean + SE = $1274.51 + 218.87$ ms, $N = 50$; two-way mixed factorial ANOVA: $F_{1,48} = 0.004$, $P = 0.95$), nor any significant differences in total gaze duration towards the man (mean + SE =

$960.00 + 192.75$ ms, $N = 50$) or woman (mean + SE = $1291.76 + 279.04$ ms, $N = 50$; two-way mixed factorial ANOVA: $F_{1,48} = 0.004$, $P = 0.95$). There was also no interaction between the orientation of the assistants (the man on the right or the woman on the right) and

side on the total gaze duration (two-way mixed factorial ANOVA: $F_{1,48} = 2.01$, $P = 0.16$).

Gaze Responses Following Playback

Analysis of the two 'correct matching' scores showed that, overall, the proportion of dogs responding correctly did not differ significantly from the expected 50% chance (binomial test: First look correct score (40%): $N = 50$, $P = 0.20$; Total gaze correct score: (50%): $N = 50$, $P = 1.00$).

However, the binary logistic regression analyses revealed a significant positive correlation between the number of adults living with the subject (household composition) and the proportion of correct responses for both the Total gaze correct scores (binary logistic regression: $\text{Wald}_1 = 7.76$, $P < 0.01$) and the First look correct scores (binary logistic regression: $\text{Wald}_1 = 7.36$, $P < 0.01$; Fig. 3). There was also a significant interaction between household composition and the side of the correct assistant on the proportion of correct responses for the Total gaze correct scores (binary logistic regression: $\text{Wald}_1 = 4.45$, $P < 0.05$) and the First look correct scores (binary logistic regression: $\text{Wald}_1 = 7.85$, $P < 0.01$). There was no effect of the subject's age or sex, average number of hours spent with people, difference in height between the assistants, gender of the playback voice, test location, use of headphones or any of the other interactions entered on the proportion of correct responses for either response variable and these IVs were not included in the final models. Together, household composition and the interaction between this variable and the side of the correct assistant accounted for 24% of the variation in the Total gaze correct scores and 35% of the total variation in the First look correct scores (Cox and Snell R^2). Equivalent results were obtained for both response variables when the same analyses were run using the categorical version of household composition (HC Group), as there was a significant main effect of HC Group (binary logistic regression: First look correct score: $\text{Wald}_1 = 6.56$, $P < 0.05$; Total gaze correct score: $\text{Wald}_1 = 9.14$, $P < 0.01$) and a significant interaction between HC Group and the side of the correct assistant (binary logistic regression: First look correct score: $\text{Wald}_1 = 5.78$, $P < 0.05$; Total gaze correct score: $\text{Wald}_1 = 6.17$, $P < 0.05$) on the proportion of correct responses.

Planned comparisons showed that dogs living with one adult or one man and one woman (HC Group 1) performed at chance level in their Total gaze correct scores (37% correct; binomial test: $N = 35$, $P = 0.18$). However, significantly more of the dogs living

Table 1

Percentage of correct responses for Total gaze correct scores depending on the household composition (HC) group and the side of the correct assistant

HC Group	Side of correct person	N	Observed correct responses (%)	Binomial test P
1 (≤ 2 adult people in household)	Right	17	53	1.00
	Left	18	22	<0.05
2 (≥ 3 adult people in household)	Right	6	100	<0.05
	Left	9	67	0.51

with three to five adults (HC Group 2) responded correctly (Fisher's exact test: $N = 15$, $P < 0.05$), which was also significantly more correct responses than expected by chance (80% correct, binomial test: $N = 15$, $P < 0.05$). Analysis of the First look correct scores revealed that subjects in HC Group 1 performed significantly below chance (29% correct; binomial test: $N = 35$, $P < 0.05$). Although HC Group 2 performed significantly better than HC Group 1 (Fisher's exact test: $N = 15$, $P < 0.05$), here their performance was not significantly above chance (67% correct; binomial test: $N = 15$, $P = 0.30$).

Separate binomial tests were then conducted for each HC Group depending on which side was correct. The side of the correct assistant was evenly distributed across both groups. For both variables, when the correct person was stood on the left side of the subject, HC Group 1 performed significantly below chance, while HC Group 2 performed at chance level. In contrast, when the correct person was stood on the right, HC Group 1 performed at chance level, while HC Group 2 performed significantly above chance (Tables 1 and 2).

Effect of Anxiety/Gaze Aversion on Observed Results

There was a significant main effect of HC Group on AB scores (two-way mixed factorial ANOVA: $F_{1,46} = 5.11$, $P < 0.05$). Dogs in HC Group 1 had significantly higher AB scores (mean + SE = $2 + 0.34$, $N = 35$) than dogs in HC Group 2 (mean + SE = $1 + 0.27$, $N = 15$). There was no significant main effect of Time (two-way mixed factorial ANOVA: $F_{1,46} = 0.31$, $P = 0.58$) or Side (two-way mixed factorial ANOVA: $F_{1,46} = 1.01$, $P = 0.32$). None of the interaction terms were significant.

The total duration of time spent looking elsewhere (at the loudspeaker and away from both assistants) after the playback presentation was not significantly different (independent measures t test: $t_{48} = 0.53$, $P = 0.60$) between dogs in HC Group 1 (mean + SE = $10922.00 + 522.19$ ms, $N = 35$) and dogs in HC Group 2 (mean + SE = $11402.67 + 675.75$ ms, $N = 15$). There was also no difference between the first response latencies (towards either assistant; independent measures t test: $t_{48} = 0.77$, $P = 0.45$) of dogs in HC Group 1 (mean + SE = $1650.00 + 456.77$ ms, $N = 35$) and dogs in HC Group 2 (mean + SE = $1090.00 + 303.25$ ms, $N = 15$). However, the amount of time spent looking away from the correct assistant after the presentation of the playback voice (i.e. time looking at either the loudspeaker, away from both assistants or at the incorrect assistant) did differ significantly between HC Groups (independent measures t test: $t_{48} = 2.05$, $P < 0.05$): dogs in HC Group 1 spent more time looking away from the correct assistant

Table 2

Percentage of correct responses for First look correct scores depending on household composition (HC) group and the side of the correct person

HC Group	Side of correct person	N	Observed correct responses (%)	Binomial test P
1 (≤ 2 adult people in household)	Right	17	47	1.00
	Left	18	11	<0.01
2 (≥ 3 adult people in household)	Right	6	100	<0.05
	Left	9	44	1.00

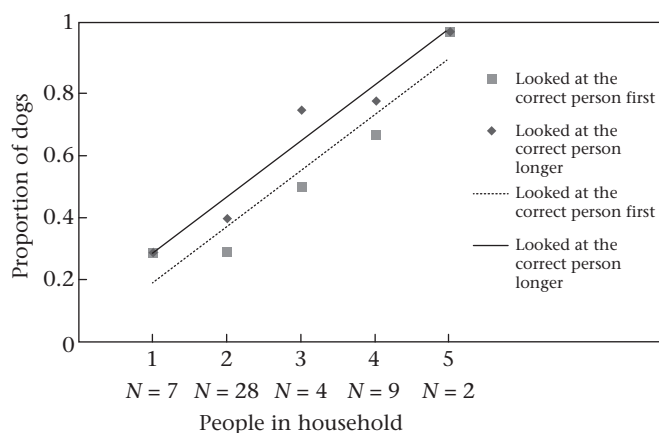


Figure 3. The proportion of dogs that looked at the correct person first and for longer depending on the number of adult people in their household. N refers to the number of dogs per group.

(mean + SE = 13 750.29 + 260.79 ms, $N = 35$) than dogs in HC Group 2 (mean + SE = 12 570.00 + 530.05 ms, $N = 15$). Similarly, dogs in HC Group 1 had significantly slower response latencies to the correct assistant (mean + SE = 7641.43 + 1000.96 ms, $N = 35$) than dogs in HC Group 2 (mean + SE = 4965.87 + 1282.18 ms, $N = 15$; independent measures t test: $t_{48} = 2.81$, $P < 0.01$). Of the total 50 subjects, 10 dogs (29%) in HC Group 1 did not look at the correct assistant at all after the presentation of the playback voice, compared to only two dogs (13%) in Group 2. Therefore, rather than looking away from both assistants, dogs in HC Group 1 spent less time looking at the correct assistant and more time looking at the incorrect assistant than dogs in HC Group 2.

DISCUSSION

Our results showed that dogs living with more people were significantly more likely to look towards a person of the same gender after hearing an unfamiliar human voice. Significantly more of the dogs living with three or more adults (including at least one man and woman) looked first, and for longer, towards the unfamiliar person that matched the gender of the voice than at the nonmatching person. Significantly more of these dogs (80%) also looked for longer at the correct person than was expected by chance. Conversely, a significantly larger proportion of the dogs living with one or two adult people (71%) looked at the incorrect person first than was expected by chance. Overall performance was not the result of a general response preference to either men or women, as the gender of the playback voice did not influence the number of correct responses. There was also no effect of the subject's age or sex on the proportion of correct responses. Finally, the average amount of time per day the subject spent with people was also not found to predict performance, suggesting that regular exposure to a wider variety of people influenced responses more strongly than the quantity of time spent in human company. However, significant interactions were found between the number of people living with the subject and the side on which the matching person was stood for both the scores for the direction of the first look and the total gaze duration. If the correct person was stood on the right side of the subject, dogs living with three or more adult people performed significantly above chance, while those living with one or two adults performed at chance level. However, if the correct person was stood on their left side, dogs living with three or more adult people performed at chance level, while those living with one or two adults performed significantly below chance.

Other studies have also found generally higher performance levels in audiovisual matching tasks when stimuli are viewed on the right side (rhesus macaque, *Macaca mulatta*: Gaffan & Harrison, 1991; bottlenose dolphin, *Tursiops truncatus*: Delfour & Marten, 2006; domestic horse, *Equus caballus*: Proops & McComb, 2012). In these cases it is thought that the left hemisphere may be more strongly recruited for 'matching with sample' tasks and in identifying familiar stimuli (Rogers, 1997; Vallortigara et al., 2008). Furthermore, findings that left hemispheric lateralization is evident in responses to conspecific or familiar vocalizations in a variety of species (e.g. domestic dog: Siniscalchi, Quaranta, & Rogers, 2008; California sea lion, *Zalophus californianus*: Böye, Güntürkün, & Vauclair, 2005; rhesus macaque: Hauser & Andersson, 1994) are consistent with the potential influence of the playback voice on orientation biases. In the current study it is apparent that although this general response bias towards the right side occurred across the subjects, the effect of this bias on their performance differed according to their level of regular exposure to people. We suggest that dogs' previous social experiences with people differentially affect how they respond to an unfamiliar person. Dogs living with a larger number of people tended to look towards the correct person,

as shown by their accurate performance if the correct person was stood on their right, but were also more likely to look towards the person on their right side overall, reducing their performance to chance level if the correct person was stood on their left side. In contrast, dogs living with fewer people appeared to avoid looking at the correct person, as evidenced by their below chance level performance when the person was stood on their left, but also showed a right side response bias, resulting in their performance at chance level when the correct person was stood on the right side.

This interpretation is supported by our finding that after the presentation of the playback voice, dogs living with fewer people looked at the correct person for significantly less time than those living with more people, by instead looking more quickly and for longer towards the incorrect person. Gaze avoidance is a coping mechanism used in stressful social situations by both humans and animals (Koolhaas et al., 1999; Thompson & Waltz, 2010). Because direct eye contact is maintained during dominance displays, dogs use gaze aversion as an appeasement signal to prevent conflict during social interactions with other dogs (Bradshaw & Nott, 1995), and have been shown to avoid making eye contact when approached in a threatening manner by an unfamiliar person (Györi, Gácsi, & Miklósi, 2010; Vas, Topál, Gácsi, Miklósi, & Csányi, 2005). The fact that, during our study, the dogs living with fewer people also produced significantly more appeasement signals (e.g. licking the muzzle, yawning) than those living with more people suggests that these dogs had greater levels of social anxiety (Rugaas, 2005). Although attempts were made so that the people in the current study did not appear threatening (they did not move or look at the subject), dogs are generally more likely to show a combination of appeasement and defensive behaviour, including gaze aversion, towards unfamiliar people (Rappolt, John, & Thompson, 1979). Shelter-housed dogs, which have had less social experience with people, are also more likely to show fear–appeasement behaviour in response to unfamiliar people than dogs with more experience (Barrera, Jakovcevic, Elgier, Mustaca, & Bentosela, 2010). Although human-directed gaze in dogs is strongly affected by previous reinforcement (Bentosela, Barrera, Jakovcevic, Elgier, & Mustaca, 2008), sociability also plays an important role in looking towards unfamiliar people, as after receiving positive reinforcement training for gazing at the experimenter's face, dogs scoring higher in their level of sociability towards an unfamiliar person gazed for significantly longer at the experimenter's face during extinction trials, when the behaviour was no longer reinforced (Jakovcevic, Mustaca, & Bentosela, 2012). Therefore, we suggest that dogs living in smaller households were more socially anxious during the study, and therefore were more likely to direct their gaze away from the more salient person who they perceived was speaking.

Zangenehpour, Ghazanfar, Lewkowicz, and Zatorre (2009) also found this 'reverse effect' in a similar cross-modal paradigm with vervet monkeys, *Cercopithecus aethiops*. Subjects were presented with two videos of rhesus monkeys producing different call types and heard vocalizations matching one of the videos. The vervet monkeys indicated their ability to match the visual and auditory information by looking for significantly longer at the incorrect image, and this result was attributed to gaze aversion due to the higher negative emotional salience of the correct image. This concurs with our interpretation of our findings in the current study, and further demonstrates that the perceived emotional salience of a stimulus can result in significant differences in responses during cognitive tasks. Thus our findings stress the importance of accounting for individual life history when investigating cognitive skills in nonhuman species.

Despite the strong influence of experience on expression, overall the dogs in the current study demonstrated their ability to discriminate correctly the person that matched the gender of the

playback voice. Therefore, our results provide the first demonstration that dogs do spontaneously learn to categorize unfamiliar people as male or female, by associating cues across different sensory modalities. Importantly, these categories are clearly not dependent on any perceptual similarities between cues (as they originate from distinct modalities), which can often explain how animals generalize learning across different stimuli in training paradigms (see Zentall, Wasserman, Lazareva, Thompson, and Rattermann (2008) for a review). However, as multiple cues were made available to the dogs in each sensory modality, we cannot yet determine the specific cues that the dogs associated across modalities, and if these cues are the same as those used by humans to categorize gender. As dogs may rely on different cues to humans when associating objects (Van der Zee, Zulch, & Mills, 2012), it could be that dogs also use different information to categorize human gender. For example, although dogs are perceptually aware of anatomically derived gender-specific cues (formant dispersion and F0) in human voices (Baru, 1975), men and women also use different intonation patterns when speaking to dogs (Prato-Previde et al., 2005). This was also seen in the current study, as the F0 range in the female voices was larger than the male range. Therefore gender differences in intonation could provide an alternative means for dogs to discriminate the gender of human voices than through the use of anatomically derived vocal differences.

Alternatively, it may be that dogs recognize more abstract correspondences between voices and people, such as matching a low-pitched voice to a person with a larger body size. Although the difference in height between the man and woman was not found to influence responses in the current study, the heights were bimodally distributed, and body weight may have also been a contributing factor. Dogs can match growls to conspecifics according to their body size (Taylor et al., 2011); thus we cannot discount the possibility that dogs also match voices to people based on body size differences rather than gender-specific cues. Further research is therefore necessary to determine more precisely how dogs learn to categorize human gender, and thus the extent to which this functionally equivalent category is comparable to the way that we categorize human gender.

Although the bases of this ability remain to be established, our observations suggest that dogs can categorize human gender in both visual/olfactory and auditory modalities. This is consistent with reports that dogs behave differently towards unfamiliar people depending on their gender, often by responding more negatively towards men (Hennessy et al., 1998; Lore & Eisenberg, 1986; Wells & Hepper, 1999), including biting men significantly more often (e.g. Rosado, García-Belenguer, León, & Palacio, 2009). While there does not appear to be gender differences in owner attachment levels towards dogs (Prato-Previde et al., 2005), male and female owners do differ in their interaction style with dogs, as men speak to their dogs less frequently (Prato-Previde et al., 2005) and are less likely to perceive their dog as being stressed (Mariti et al., 2012). It is possible that gender-specific behavioural differences may create a need for dogs to categorize men and women in order to adapt their responses appropriately. Determining more specifically how and why dogs learn to categorize men and women has important practical implications for understanding their responses to different people. While our study has demonstrated that multi-sensory cues, including vocal cues, are associated by dogs, we have yet to determine which specific cues are used and how these may influence responses.

Conclusion

Our findings illustrate that dogs can spontaneously categorize human gender by associating cues across sensory modalities. The

strong influence of the dogs' social experience with humans in the expression of this ability also highlights the important issue of accounting for life history as a source of individual variation in the natural expression of cognitive abilities by nonhuman species. Investigating how animals perceive and categorize their social environment is a crucial step towards understanding the nature of interactions between domesticated animals and humans.

Acknowledgments

We thank all of the owners who allowed their dogs to participate in our study. We are very grateful to The Dog Hut and Hamsey Riding School for the use of their facilities during testing. We also thank James Day, Sarah Jeannin, Olessya Zhuravleva, Mike Fitzgerald, Mariam Zia, Chris Girvan, Elmi Terjavainena, Jennifer Wathan and Mahmoud Elsherif for their assistance during the data collection and Solène Derville and Lotte Carlebach for second coding the video data.

References

- Adachi, I., Kuwahata, H., & Fujita, K. (2007). Dogs recall their owner's face upon hearing the owner's voice. *Animal Cognition*, 10(1), 17–21. <http://dx.doi.org/10.1007/s10071-006-0025-8>.
- Autier-Dérian, D., Deputte, B. L., Chalvet-Monfray, K., Coulon, M., & Mounier, L. (2013). Visual discrimination of species in dogs (*Canis familiaris*). *Animal Cognition*, 16, 637–651. <http://dx.doi.org/10.1007/s10071-013-0600-8>.
- Bachorowski, J. A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *Journal of the Acoustical Society of America*, 106, 1054–1063. <http://dx.doi.org/10.1121/1.427115>.
- Barrera, G., Jakovcovic, A., Elgier, A. M., Mustaca, A., & Bentosela, M. (2010). Responses of shelter and family dogs to an unknown human. *Journal of Veterinary Behavior: Clinical Applications and Research*, 5(6), 339–344. <http://dx.doi.org/10.1016/j.jveb.2010.08.012>.
- Baru, A. V. (1975). Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration and number of formants) in dog. In G. Fant, & M. A. A. Tatham (Eds.), *Auditory analysis and perception of speech* (pp. 91–101). New York: Academic Press.
- Bentosela, M., Barrera, G., Jakovcovic, A., Elgier, A. M., & Mustaca, A. E. (2008). Effect of reinforcement, reinforcer omission and extinction on a communicative response in domestic dogs, (*Canis familiaris*). *Behavioural Processes*, 78, 464–469. <http://dx.doi.org/10.1016/j.beproc.2008.03.004>.
- Böye, M., Güntürkün, O., & Vauclair, J. (2005). Right ear advantage for conspecific calls in adults and subadults, but not infants, California sea lions (*Zalophus californianus*): hemispheric specialization for communication? *European Journal of Neuroscience*, 21, 1727–1732. <http://dx.doi.org/10.1111/j.1460-9568.2005.04005.x>.
- Bradshaw, J. W. S., & Nott, H. M. R. (1995). Social and communication behaviour of companion dogs. In J. Serpell (Ed.), *The domestic dog: Its evolution, behaviour and interactions with people* (pp. 115–130). Cambridge, U.K.: Cambridge University Press.
- Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making inferences about the location of hidden food: social dog, causal ape. *Journal of Comparative Psychology*, 120(1), 38–47. <http://dx.doi.org/10.1037/0735-7036.120.1.38>.
- Burton, M. A., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, 22(2), 153–176. <http://dx.doi.org/10.1068/p220153>.
- Delfour, F., & Marten, K. (2006). Lateralized visual behavior in bottlenose dolphins (*Tursiops truncatus*) performing audio-visual tasks: the right visual field advantage. *Behavioural Processes*, 71(1), 41–50. <http://dx.doi.org/10.1016/j.beproc.2005.09.005>.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague, The Netherlands: Mouton.
- Faragó, T., Pongrácz, P., Miklósi, Á., Huber, L., Virányi, Z., & Range, F. (2010). Dogs' expectation about signalers' body size by virtue of their growls. *PLoS One*, 5(12), e15175. <http://dx.doi.org/10.1371/journal.pone.0015175>.
- Gaffan, D., & Harrison, S. (1991). Auditory-visual associations, hemispheric specialization and temporal-frontal interaction in the rhesus monkey. *Brain*, 114(5), 2133–2144. <http://dx.doi.org/10.1093/brain/114.5.2133>.
- Goldstein, U. G. (1980). *An articulatory model for the vocal tracts of growing children* (Unpublished doctoral dissertation). Retrieved from: Massachusetts Institute of Technology <http://hdl.handle.net/1721.1/16118>.
- Györi, B., Gácsi, M., & Miklósi, A. (2010). Friend or foe: context dependent sensitivity to human behaviour in dogs. *Applied Animal Behaviour Science*, 128, 69–77. <http://dx.doi.org/10.1016/j.applanim.2010.10.005>.

- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of cognition in dogs. *Science*, 298(5598), 1634–1636. <http://dx.doi.org/10.1126/science.1072702>.
- Hauser, M. D., & Andersson, K. (1994). Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proceedings of the National Academy of Sciences*, 91(9), 3946–3948. Retrieved from: <http://www.pnas.org/content/91/9/3946.full.pdf+html>.
- Heffner, H. (1975). Perception of biologically meaningful sounds by dogs. *Journal of the Acoustical Society of America*, 58, S124. <http://dx.doi.org/10.1121/1.2001931>.
- Hennessy, M. B., Williams, M. T., Miller, D. D., Douglas, C. W., & Voith, V. L. (1998). Influence of male and female petters on plasma cortisol and behaviour: can human interaction reduce the stress of dogs in a public animal shelter? *Applied Animal Behaviour Science*, 61, 63–77. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0168159198001798#>.
- Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: the role of shape and colour. *Proceedings of the Royal Society B: Biological Sciences*, 261, 367–373. <http://dx.doi.org/10.1098/rspb.1995.0161>.
- Jakovcic, A., Mustaca, A., & Bentosela, M. (2012). Do more sociable dogs gaze longer to the human face than less sociable ones? *Behavioural Processes*, 90, 217–222. <http://dx.doi.org/10.1016/j.beproc.2012.01.010>.
- Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., et al. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, 23, 925–935. [http://dx.doi.org/10.1016/S0149-7634\(99\)00026-3](http://dx.doi.org/10.1016/S0149-7634(99)00026-3).
- Lore, R. K., & Eisenberg, F. B. (1986). Avoidance reactions of domestic dogs to unfamiliar male and female humans in a kennel setting. *Applied Animal Behaviour Science*, 15(3), 262–266. [http://dx.doi.org/10.1016/0168-1591\(86\)90096-1](http://dx.doi.org/10.1016/0168-1591(86)90096-1).
- Mariti, C., Gazzano, A., Moore, J. L., Baragli, P., Chelli, L., & Sigheiri, C. (2012). Perception of dogs' stress by their owners. *Journal of Veterinary Behavior: Clinical Applications and Research*, 7(4), 213–219. <http://dx.doi.org/10.1016/j.jveb.2011.09.004>.
- Miller, P. E., & Murphy, C. J. (1995). Vision in dogs. *Journal of the American Veterinary Medical Association*, 207(12), 1623–1634. Retrieved from: <http://redwood.berkeley.edu/bruno/animal-eyes/dog-vision-miller-murphy.pdf>.
- Murai, C., & Tomonaga, M., Kamegai, K., Terazawa, N., & Yamaguchi, M. K. (2004). Do infant Japanese macaques (*Macaca fuscata*) categorize objects without specific training? *Primates*, 45(1), 1–6. <http://dx.doi.org/10.1007/s10329-003-0053-9>.
- Murai, C., Kosugi, D., Tomonaga, M., Tanaka, M., Matsuzawa, T., & Itakura, S. (2005). Can chimpanzee infants (*Pan troglodytes*) form categorical representations in the same manner as human infants (*Homo sapiens*)? *Developmental Science*, 8(3), 240–254. <http://dx.doi.org/10.1111/j.1467-7687.2005.00413.x>.
- O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D., & Abdi, H. (1998). The perception of face gender: the role of stimulus structure in recognition and classification. *Memory & Cognition*, 26(1), 146–160. <http://dx.doi.org/10.3758/BF03211378>.
- Prato-Previde, E., Fallani, G., & Valsecchi, P. (2005). Gender differences in owners interacting with pet dogs: an observational study. *Ethology*, 112(1), 64–73. <http://dx.doi.org/10.1111/j.1439-0310.2006.01123.x>.
- Proops, L., & McComb, K. (2012). Cross-modal individual recognition in domestic horses (*Equus caballus*) extends to familiar humans. *Proceedings of the Royal Society B: Biological Sciences*, 279, 3131–3138. <http://dx.doi.org/10.1098/rspb.2012.0626>.
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition*, 11(2), 339–347. <http://dx.doi.org/10.1007/s10071-007-0123-2>.
- Rappolt, G. A., John, J., & Thompson, N. S. (1979). Canine responses to familiar and unfamiliar humans. *Aggressive Behavior*, 5(2), 155–161. [http://dx.doi.org/10.1002/1098-2337\(1979\)5:2<155::AID-AB2480050206>3.0.CO;2-D](http://dx.doi.org/10.1002/1098-2337(1979)5:2<155::AID-AB2480050206>3.0.CO;2-D).
- Rogers, L. J. (1997). Early experiential effects on laterality: research on chicks has relevance to other species. *Laterality*, 2(3–4), 199–219. <http://dx.doi.org/10.1080/713754277>.
- Rosado, B., García-Belenguer, S., León, M., & Palacio, J. (2009). A comprehensive study of dog bites in Spain, 1995–2004. *Veterinary Journal*, 179(3), 383–391. <http://dx.doi.org/10.1016/j.tvjl.2008.02.002>.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439. [http://dx.doi.org/10.1016/0010-0285\(76\)90013-X](http://dx.doi.org/10.1016/0010-0285(76)90013-X).
- Rugaas, T. (2005). *On talking terms with dogs: Calming signals*. Wenatchee, WA: Dogwise Publishing.
- Savolainen, P., Zhang, Y., Luo, J., Lundeberg, J., & Leitner, T. (2002). Genetic evidence for an East Asian origin of domestic dogs. *Science*, 298, 1610–1613. <http://dx.doi.org/10.1126/science.1073906>.
- Siniscalchi, M., Quaranta, A., & Rogers, L. J. (2008). Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS One*, 3(10), e3349. <http://dx.doi.org/10.1371/journal.pone.0003349>.
- Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: a case study for colour. *Behavioral and Brain Sciences*, 28(04), 469–529. <http://dx.doi.org/10.1017/S0140525X05270085>.
- Taylor, A. M., Reby, D., & McComb, K. (2011). Cross modal perception of body size in domestic dogs (*Canis familiaris*). *PLoS One*, 6, 1–6. <http://dx.doi.org/10.1371/journal.pone.0017069>.
- Thompson, B. L., & Waltz, J. (2010). Mindfulness and experiential avoidance as predictors of posttraumatic stress disorder avoidance symptom severity. *Journal of Anxiety Disorders*, 24(4), 409–415. <http://dx.doi.org/10.1016/j.janxdis.2010.02.005>.
- Titze, I. R. (2000). *Principles of voice production*. Iowa City, IA: National Center for Voice and Speech.
- Vallortigara, G., Snyder, A., Kaplan, G., Bateson, P., Clayton, N. S., & Rogers, L. J. (2008). Are animals autistic savants? *PLoS Biology*, 6, 208–214. <http://dx.doi.org/10.1371/journal.pbio.0060042>.
- Van der Zee, E., Zulch, H., & Mills, D. (2012). Word generalization by a dog (*Canis familiaris*): is shape important. *PLoS One*, 7, e49382. <http://dx.doi.org/10.1371/journal.pone.0049382>.
- Vas, J., Topál, J., Gácsi, M., Miklósi, A., & Csányi, V. (2005). A friend or an enemy? Dogs' reaction to an unfamiliar person showing behavioural cues of threat and friendliness at different times. *Applied Animal Behaviour Science*, 94(1–2), 99–115. <http://dx.doi.org/10.1016/j.applanim.2005.02.001>.
- Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., & Kent, R. D. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: an imaging study. *Journal of the Acoustical Society of America*, 125, 1666–1678. <http://dx.doi.org/10.1121/1.3075589>.
- Walker-Andrews, A. S., Bahrnick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology*, 3(2), 55–75. http://dx.doi.org/10.1207/s15326969eco0302_1.
- Wells, D. L., & Hepper, P. G. (1999). Male and female dogs respond differently to men and women. *Applied Animal Behaviour Science*, 61(4), 341–349. [http://dx.doi.org/10.1016/S0168-1591\(98\)00202-0](http://dx.doi.org/10.1016/S0168-1591(98)00202-0).
- Zangenehpour, S., Ghazanfar, A. A., Lewkowicz, D. J., & Zatorre, R. J. (2009). Heterochrony and cross-species intersensory matching by infant vervet monkeys. *PLoS One*, 4, e4302. <http://dx.doi.org/10.1371/journal.pone.0004302>.
- Zentall, T. R., Wasserman, E. A., Lazareva, O. F., Thompson, R. K. R., & Rattermann, M. J. (2008). Concept learning in animals. *Comparative Cognition and Behavior Reviews*, 3, 13–45. <http://dx.doi.org/10.3819/ccbr.2008.30002>.